On Nonlinear Modeling of Microwave Devices Using Interpolating Wavelets

Mikhail Toupikov, *Member, IEEE*, Guangwen (George) Pan, *Senior Member, IEEE*, and Barry K. Gilbert, *Fellow, IEEE*

***Abstract*—Nonlinear semiconductor devices are modeled using the sparse point representation based upon interpolating wavelets. The functions of potential, fields, electron, and hole current densities inside the device are represented by a twofold expansion in scaling functions and wavelets. In most regions where the functions are smoothly varying, only scaling functions are employed as the bases. In contrast, in small regions with sharp material or field variations, additional basis functions, i.e., wavelets, are introduced. A nonuniform mesh generated in this manner is fully adaptive, dynamic, and object oriented. Examples of device simulations are presented, demonstrating good agreement with published literature and commercial software. The numerical examples also show substantial savings in computer memory for electrically large problems.**

***Index Terms*—Interpolating wavelets, microwave devices, nonlinear modeling.**

I. Introduction

DENSELY packed modern integrated circuits consist of closely spaced active and passive devices, with many levels of interconnect lines and discontinuities. The highest speed digital integrated circuits currently under development are employing completely new transistors such as silicon germanium (SiGe), gallium arsenide (GaAs), and indium phosphide (InP) heterojunction bipolar transistors (HBT's), with upper cutoff frequencies of 45, 60, and even 225 GHz. Even CMOS-on-insulator (CMOS/SOI) integrated circuits are capable of operating at clock rates in the 500–2000-MHz range, while the SiGe, GaAs, and InP chips have been demonstrated to operate at clock rates as high as 10–50 GHz. The interconnects on these high-performance components must increasingly be treated as transmission lines rather than merely as distributed capacitive loads. In such cases, the performance of an entire integrated circuit may be adversely affected by the high device packing density and/or high-speed operation, due to unwanted effects such as, but not limited to, crosstalk caused by coupling, surface waves, and unintended radiation. It is thus apparent that integrated circuits designed with these transistors must be developed based on advanced design tools that consider all the circuit elements simultaneously, including active devices, passive components, radiation elements, and the electronic packaging into which the integrated circuits are placed. The possibility of achieving this type of global circuit modeling approach has been demonstrated in [1] and [2].

Recently, a new category of orthogonal systems, referred to as "orthogonal wavelets," has appeared in the literature [3]. The wavelet expansion technique has proven to be an efficient method in the approximation of functions. Different classes of wavelets have been used to analyze a variety of passive circuit components. For example, Daubechies wavelets have been employed in [4] for the analysis of microstrip floating structures. In [5], Battle–Lemarie wavelets have been imposed for the extraction of the internal inductance and skin-effect resistance matrices of multiple lossy transmission lines. In this paper, we shall consider the application of multiresolution analysis to the modeling of active semiconductor devices with nonlinearities.

The use of scaling functions and wavelets as a complete set of basis functions is referred to as multiresolution analysis [6]. To derive a new algorithm, the potential distribution inside the semiconductor and electron and hole current densities are represented by a twofold expansion in scaling functions and wavelets. Using only scaling functions allows correct modeling of smoothly varying electromagnetic fields and materials parameters. In regions with large field variations, additional basis functions (wavelets) are introduced. In our derivations, we have employed a special class of wavelets; namely, the interpolating wavelets. This wavelet system has already been applied to the solution of boundary problems for partial differential equations (PDE's). For this type of wavelet, the evaluation of differential operators is simplified due to their elegant representations in terms of cubic polynomial functions in the spatial domain.

Several different approaches for solving PDE's using wavelets have been proposed, including the multiresolution time-domain (MRTD) methods with adaptive gridding [7], [8]. In [9], the author used wavelets for finding regions requiring grid refinement in a finite-difference method. It has been noted by several authors that nonlinear operators such as multiplication are too computationally expensive when performed directly in a wavelet basis, and several attempts have been made to deal with this problem. Keiser [10] has used Coifman wavelets to obtain approximations of point values in a wavelet method, thus simplifying the treatment of nonlinearities. In this paper, we will explore and extend an approach proposed in [11] to deal with nonlinearities using the so-called sparse point representation (SPR).

Modeling of nonlinear semiconductor devices, e.g., transistors or diodes, produces functions (carrier concentration and potential distribution) that are smooth almost everywhere in the

Manuscript received October 19, 1998. This work was supported in part by the Microsystems Technology Office of the Defense Advanced Research Projects Agency under Contract N66001-94-C-0051 from the Space and Warfare Systems Center, San Diego, CA.

M. Toupikov and G. Pan are with the Department of Electrical Engineering, Arizona State University, Tempe, AZ 85287-5706 USA.

B. K. Gilbert is with the Department of Physiology and Biophysics, Mayo Foundation, Rochester, MN 55905-0001 USA.

Publisher Item Identifier S 0018-9480(00)02785-X.

domain, except in a small region of sharp variation near the p-n junction. It is ideal to have a nonuniform grid, which is fine in the region around the sharp variation and coarse in areas where the solution is smooth. As will be described in the following sections, the SPR can provide a nonuniform grid that is fully adaptive and dynamic. In other words, the SPR is fully object oriented. The SPR itself is based on the new concept of interpolating wavelets. The application of the SPR creates an opportunity to consider the grid as a dynamic object that is fully integrated into the solution. The nonuniform grid becomes fully adaptive, in that changes in the grid follow the changes in the solution for each time step.

II. Interpolating Subdivision Scheme

Introduced by Deslauriers and Dubuc [12], the dyadic grids on the real line (or the subspace of the scaling functions)

$$V_j = \left\{ x_{j,k} \in \mathcal{R} | x_{j,k} = 2^{-j}k,\, k \in \mathcal{Z} \right\}, \qquad j \in \mathcal{Z} \tag{1}$$

i.e., the grid points, are the integers in V_0 and half-integers for V_1. In general, the dyadic grid V_{j+1} contains all the grid points in V_j, as well as additional points inserted halfway in between each of the points in V_j. More information and additional references describing interpolating subdivision schemes can be found in [3].

Given function values on V_j, $\{f_{j,k}\}_{k\in\mathcal{Z}}$, where $f_{j,k} = f(x_{j,k})$ is a function defined on the grid points in V_j, the interpolating subdivision scheme defines $f_{j+1,k}$ in V_{j+1}. The even numbered grid points $x_{j+1,2k}$ already exist in V_j, and the corresponding function values are left unchanged. Values at the odd grid points $x_{j+1,2k+1}$ are computed by polynomial interpolation from the values at the even grid points. We denote this interpolating polynomial by $P_{j+1,2k+1}$. The degree $p-1$ of this polynomial is odd to make the scheme symmetric, i.e., we interpolate from an even number of function values. Formally, we define one step of the subdivision scheme as

$$\begin{cases} f_{j+1,2k} = f_{j,k} \\ f_{j+1,2k+1} = P_{j+1,2k+1}(x_{j+1,2k+1}) \end{cases} \qquad \forall k \in \mathcal{Z} \tag{2}$$

where $P_{j+1,2k+1}(x)$ is chosen such that

$$P_{j+1,2k+1}(x_{j,k+l}) = f_{j,k+l}, \qquad \text{for } -p/2 < l \le p/2. \tag{3}$$

Thus, we use p symmetric points on the coarser grid V_j to interpolate one new function value on the finer grid V_{j+1}. For dyadic grids, we can explicitly define the interpolating polynomial. For the case $p = 4$, a cubic polynomial, the computed values at odd grid points are

$$f_{j+1,2k+1} = \left(-f_{j,k-1} + 9f_{j,k} + 9f_{j,k+1} - f_{j,k+2} \right) \Big/ 16.$$

Repeating the aforementioned subdivision recursively, we obtain representations on successively finer grids V_j as j increases, and in the limit $j \to \infty$, we have a representation of the function $f(x)$ at all dyadic rational points.

If the subdivision starts with the Kronecker delta sequence $\{\delta_{0,k}\}_{k\in\mathcal{Z}}$ on V_0 and is then refined to V_j, in the limit $j \to \infty$, we will obtain the scaling function of the interpolating wavelets $\varphi(x)$. From the construction, it follows that $\varphi(x)$ has a compact support $[-p+1, p-1]$ and is symmetric around $x = 0$. If we make one step in the subdivision scheme for the sequence $\{\delta_{0,k}\}$, we obtain the two-scale relation

$$\varphi(x) = \sum_{k=-p+1}^{p-1} \varphi(k/2)\varphi(2x-k). \tag{4}$$

Using an integer translation of $\varphi(x)$, we have a basis in V_0, and the interpolant of any continuous function $f(x)$ in V_0 can be defined as

$$\mathcal{P}f(x) = \sum_k f_{0,k}\varphi(x-k).$$

The interpolant of any continuous function $f(x)$ in V_j can be defined as

$$\mathcal{P}_j f(x) = \sum_k f_{j,k}\varphi_{j,k}(x)$$

where $\varphi_{j,k}(x) = \varphi(2^j x - k)$, $k \in \mathcal{Z}$ is a basis in V_j. Here, notation V_j is used as a function space and grid. Since the basis functions are cardinal, $\varphi_{j,k}(x_{j,l}) = \delta_{k,l}$, $j, k, l \in \mathcal{Z}$, there is a one-to-one correspondence between grid points and basis functions.

The scaling function spaces introduced above generate a ladder of spaces

$$\cdots \subset V_{j-1} \subset V_j \subset V_{j+1} \subset \cdots$$

and the interpolating scheme enables us to move through these spaces (i.e., to achieve either refinement or coarsening). Additional spaces W_j can be introduced to encode the difference between V_j and V_{j+1}

$$V_{j+1} = V_j \bigoplus W_j.$$

Introducing a basis $\{\psi_{j,k}\}_{k\in\mathcal{Z}}$ in W_j, we can write

$$\mathcal{P}_{j+1}f(x) - \mathcal{P}_j f(x) = \sum_k d_{j,k}\psi_{j,k}(x)$$

where $\psi_{j,k}(x) = \psi(2^j - k)$. The function $\psi(x)$ is a wavelet and $d_{j,k}$ are wavelet coefficients. One of the simplest possible choices is to define $\psi(x)$ as

$$\psi(x) = \varphi(2x-1).$$

This wavelet was introduced by Donoho [13]. This choice is called the Lazy wavelet transform [14]. This transform does not perform any operation other than subsampling into the odd indexed samples.

Given a representation of a function in the space V_{j+1}, one can decompose it into a coarser scale representation in V_j and a correction in W_j. Starting with a representation in V_J, this decomposition can be repeated $J - j_0$ times

$$\sum_k f_{J,k}\varphi_{J,k}(x) = \sum_k f_{j_0,k}\varphi_{j_0,k}(x) + \sum_{j_0 \le j < J} \sum_k d_{j,k}\psi_{j,k}(x).$$

On the right-hand side, our function is decomposed into the scaling function representation on a coarse grid V_{j_0} and wavelets on successively finer scales.

III. Sparse Point Representation

The idea behind the use of a wavelet basis is that certain functions are well compressed in such a basis. As a result, only a few basis functions are needed to represent the function with a small error. Assume that a function is represented by N points on a uniform grid, and the same function is represented, with an error ϵ, by N_s wavelet coefficients, where $N_s \ll N$. We would like to be able to compute derivatives and multiply functions in this wavelets basis in $O(N_s)$ time. The interpolating wavelet transform provides the means to achieve this goal. The chosen basis has the property that each wavelet coefficient corresponds to a function value at a grid point.

Assume that we have the wavelet representation

$$\mathcal{P}_J f(x) = \sum_k f_{j_0,k}\varphi_{j_0,k}(x) + \sum_{j=j_0}^{J-1}\sum_k d_{j,k}\psi_{j,k}(x).$$

Operations such as differentiation and multiplication can be costly when performed in a wavelet basis due to interactions between scales in a wavelet representation. It would be ideal to transform the N_s wavelet coefficients to N_s point values. Such a transform does exist for the interpolating wavelets due to the one-to-one correspondence between wavelet coefficients and point values. To obtain a sparse wavelet representation, we remove all wavelet coefficients with magnitude less than some threshold value ϵ. We then have the threshold expansion

$$\mathcal{P}_J f(x) = \sum_k f_{j_0,k}\varphi_{j_0,k}(x) + \sum_{(j,k)\in I(\epsilon)}\sum_k d_{j,k}\psi_{j,k}(x) \tag{5}$$

where the set $I(\epsilon)$ contains indexes of all significant coefficients. The inverse transform can be performed, but only for those points that correspond to the significant wavelet coefficients in $I(\epsilon)$. If any point value is needed that does not exist, it will be interpolated from the coarser scale recursively. The algorithm will terminate since we have all values on the coarsest grid V_{j_0}.

This inverse transform leads us to an SPR. Note that the SPR is not a representation in a basis; rather, it is simply a collection of point values $\{f_{j,k}\}_{(j,k)\in I(\epsilon)}$. The SPR can be computed without explicitly forming a sparse wavelet representation, i.e., it is possible to store the point values in the SPR, instead of the wavelet coefficients. The wavelet coefficients are only computed to decide if the corresponding point value is to be included in the SPR or not.

To examine the approximation error arising from using the threshold expansion (5), we need the maximum norm

$$|g|_\infty = \max_{0\le x\le 1}|g(x)|.$$

We are interested in the dependence of the error on the threshold parameter ϵ. Donoho [13] and Holmström [11] have shown that the estimation

$$|f(x) - \mathcal{P}_J f(x)|_\infty \le c_1\epsilon$$

holds for a sufficiently smooth function $f(x)$ and for a large enough level J. Further, the number of significant coefficients N_s depends on ϵ as

$$N_s \le c_2\epsilon^{-1/p}$$

or equivalently

$$\epsilon \le c_2^p N_s^{-p}.$$

Combining the last three inequalities, we can achieve a bound on the error versus N_s as

$$|f(x) - \mathcal{P}_J f(x)|_\infty \le c_3 N_s^{-p} \tag{6}$$

where c_i $(i = 1, 2, 3)$ denote constants for a given function $f(x)$. This result indicates that the sparse interpolating wavelet approximation is of order p in the number of significant coefficients N_s.

To perform the multiplication in $O(N_s)$ time, we need to specify the SPR pattern of the product. The SPR of the product can be chosen as the union of the two operand representations. If a point value is missing, it is again interpolated from the coarser scale in the SPR.

Differentiation can be applied to the SPR of the function. For each point for which we wish to approximate the derivative, we locate the closest point in the SPR and choose the distance to that point as the step length h. A centered finite-difference stencil of order p can then be applied, where p is the order of the interpolating wavelets in the SPR. If any point is missing, it can be interpolated from a coarser scale. If any point in the stencil is located outside the boundary, a one-sided stencil of the same order is employed. The finite-difference approximation of the first and second derivatives are, respectively,

$$f'(x) \approx \frac{1}{h}\sum_i g_i' f(x+ih)$$

and

$$f''(x) \approx \frac{1}{h^2}\sum_i g_i'' f(x+ih).$$

On an interval, the filter coefficients g_i' and g_i'' depend on x since a one-sided approximation near the boundaries is used; their values for the case $p = 4$ are presented in Tables I and II.

In these tables, the filter coefficients for the first and second derivatives are shown at the left-hand-side boundary. The coefficients at the right-hand-side boundary are reversed in order, with opposite signs. When the threshold parameter $\epsilon \to 0$, the above finite-difference approximations become ordinary finite-difference approximations on a uniform grid.

In the case of two dimensions, partial derivatives in each direction are evaluated using the one-dimensional (1-D)

TABLE I
Filter Coefficients g_i' for the First Derivative Approximation

n	-2	-1	0	1	2	3	4
$0 \le x < h$			-25/12	4	-3	4/3	-1/4
$h \le x < 2h$		-1/4	-5/6	3/2	-1/2	1/12	
$x \ge$	1/12	-2/3	0	2/3	-1/12		

TABLE II
FILTER COEFFICIENTS g_i'' FOR THE SECOND DERIVATIVE APPROXIMATION

n	-2	-1	0	1	2	3	4	5
$0 \le x < h$			15/4	-77/6	-107/6	-13	61/12	-5/6
$h \le x < 2h$		5/6	-5/4	-1/3	7/6	-1/2	1/12	
$x \ge$	-1/12	4/3	-5/2	4/3	-1/12			

approximation. The step length h is chosen as the distance to the closest point in the SPR, as measured along any of the coordinate directions.

IV. STATEMENT OF THE PROBLEM

The basic physical model consists of three coupled PDE's: the Poisson equation for the electric field and two continuity equations for the electrons and holes. These equations are supplemented by the expressions for the electron and hole current densities. Since the goal of this paper is to apply wavelets to simplify the analysis, the drift–diffusion approximation is used in the current density expressions. The mathematical model consists of Poisson's equation

$$\nabla^2 U = -\frac{q}{\epsilon}(N_d - N_a + p - n). \tag{7}$$

The electron and hole carrier concentrations are found from the continuity equations

$$\frac{\partial n}{\partial t} - \frac{1}{q}\nabla \cdot \boldsymbol{J}_n = 0 \tag{8}$$

$$\frac{\partial p}{\partial t} + \frac{1}{q}\nabla \cdot \boldsymbol{J}_p = 0 \tag{9}$$

and the electron and hole current densities satisfy

$$\boldsymbol{J}_n = -q\mu_n n \nabla U + qD_n \nabla n \tag{10}$$

$$\boldsymbol{J}_p = -q\mu_p p \nabla U - qD_p \nabla p. \tag{11}$$

The first term in (10) and (11) represents the conductivity current due to the electric field, and the second term represents a current flow due to diffusion.

Assuming the Einstein relation for both electrons and holes [15]

$$D_n = \mu_n \frac{k_B T}{q}$$
$$D_p = \mu_p \frac{k_B T}{q}.$$

In the following derivations, we consider the mobility of the carriers μ_n and μ_p to be constant and field independent. More complicated cases can be solved without major difficulties, although these cases may be more computationally expensive. Equations (7)–(11) summarize the coupled system of PDE's describing the semiconductor device. All that remains is to specify boundary conditions for a particular geometry.

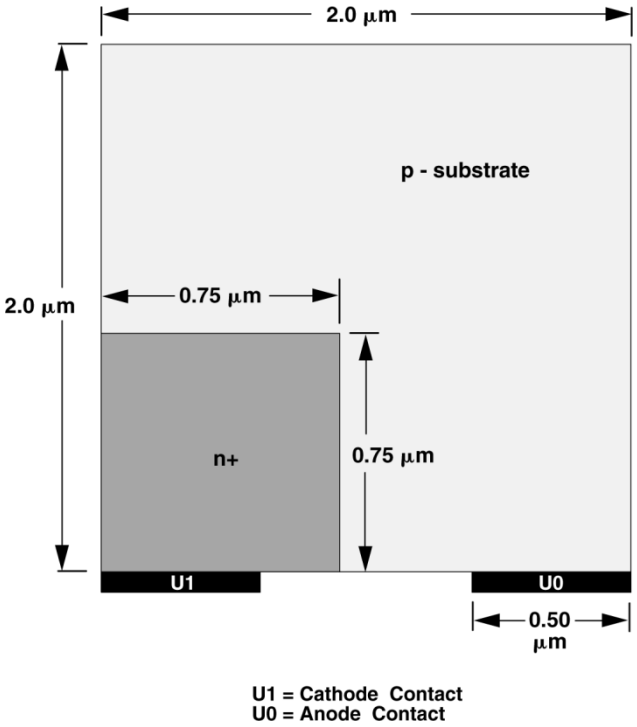


Fig. 1. Idealized 2-D cross section of a silicon diode.

Fig. 1 presents a representative example of a (2-D) cross section of a silicon abrupt diode. The potential, electron, and hole carrier concentrations satisfy appropriate initial, boundary, and interface conditions. In general, there are semiconductor/conductor interfaces (contacts), semiconductor/isolation interfaces, and external boundaries. We assume thermal equilibrium and charge neutrality on ohmic contacts

$$np = n_i^2$$

$$p - n + D = 0$$

where n_i is an intrinsic carrier concentration and D is a given doping function.

These conditions are provided by a condition on the electric potential, which is given by the built-in voltage $U_{\rm bi}$ and the applied potential U_a

$$U = U_a + U_{\rm bi}.$$

At the outside boundaries, we always assume a vanishing outward electric field and vanishing outward current densities

$$\nabla U \cdot \hat{\boldsymbol{n}} = \boldsymbol{J}_n \cdot \hat{\boldsymbol{n}} = \boldsymbol{J}_p \cdot \hat{\boldsymbol{n}} = 0.$$

In order to solve the drift–diffusion equations, it is convenient to express all variables in terms of scaled quantities. Several different scaling approaches are possible [16], [17]. In this study, we have followed the approach of De Mari [16]. Without explicit declaration, we always assume that the quantities are scaled.

The spatial discretization is performed using a method proposed by Scharfetter–Gummel [18]. The discrete model, therefore, consists of a system of two nonlinear ordinary differential equations (ODE's) of the first order and an algebraic equation. The two ODE's are coupled through the algebraic equation.

V. Transient Solution

In scaled form, the basic semiconductor equations can be written as

$$\begin{aligned} g_1(\boldsymbol{U}, \boldsymbol{n}, \boldsymbol{p}) &= 0 \\ g_2(\boldsymbol{U}, \boldsymbol{n}, \boldsymbol{p}) &= 0 \\ g_3(\boldsymbol{U}, \boldsymbol{n}, \boldsymbol{p}) &= 0 \end{aligned} \tag{12}$$

where vectors $\boldsymbol{U}, \boldsymbol{n}, \boldsymbol{p}$ are now the SPR of the normalized electrostatic potential and carrier concentrations.

In our calculations we will follow the modified Gummel iterative approach [18]. The resulting system of ODE's can be solved by an ODE solver, e.g., the fourth-order Runge–Kutta method [19].

The iterative procedure to solve the problem (12) can be presented in the following seven steps.

Step 1: Set initial values for the function p and n and fix a threshold value ϵ.
Step 2: Obtain SPR for p, n, and U.
Step 3: Solve the Poisson's equation g_1 and obtain an SPR for the potential.
Step 4: Make one step in the continuity equation g_2 for **n**.
Step 5: Repeat previous step in the continuity equation g_3 for **p**.
Step 6: Update all SPR's.
Step 7: Go to Step 2.

The basic difficulty in the solution of the transient system is the requirement that the numerical method must be unconditionally stable.

For the solution of the linear algebraic system, we use a fast variant of the Bi-CG iterative solver, called Bi-CGSTAB [20].

It is very important that the Bi-CGSTAB method does not involve any use of the transpose matrix A^T. Since we use interpolation to obtain the missing points in the finite-difference procedure, it is difficult to explicitly assemble the system matrix A. Instead, we calculate all matrix vector products directly, without forming the matrix A. The matrix vector product is treated as an operator acting on the SPR of the unknown function.

VI. Grid Adaptation and Interpolating Wavelets

Accuracy and efficiency are strongly related to the discretization of the equations and, thus, to the chosen mesh. In a standard finite-difference algorithm, a tensor product mesh is usually selected. The mesh lines continue in the regions far from the junction, where the potential is a slowly varying function and there is no need for a fine mesh.

The authors of [21] and [22] have introduced methods of mesh line termination. The Laplacian discretization in any node (i, j) at the end of a terminated line is obtained as a linear combination of potential values in the surrounding nodes. The local truncation error is of the order of a third derivative of the potential and it is used as a refinement criterion. Problems of discretization and grid adaptation have also been addressed in [23].

Several different criteria for the grid refinement process are available. One can refine the mesh by taking into account the error in the Poisson equation, or the electron/hole continuity equations, or through the use of doping concentrations. Unfortunately, there is a high likelihood that a mesh that is optimal for the potential will be insufficient for a carrier concentration and vice versa. Attempts to satisfy all criteria will lead to a very dense mesh. However, the interpolating wavelets provide a unique opportunity to overcome these difficulties. The SPR of a function contains the only points that correspond to significant wavelet coefficients.

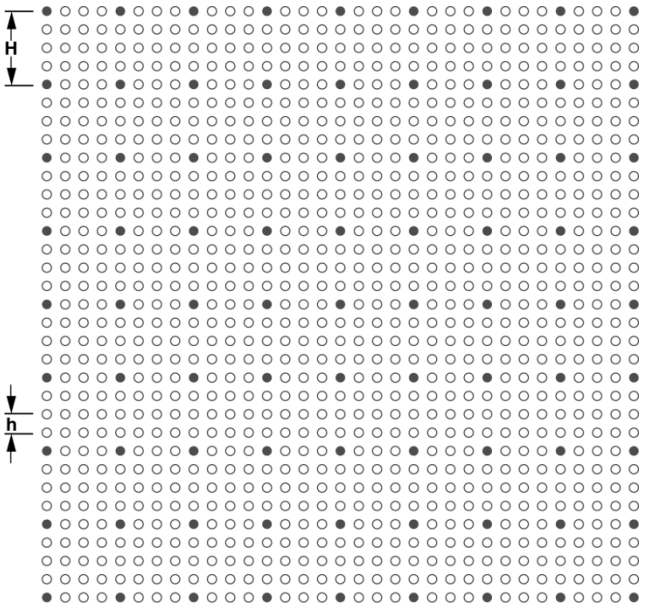


Fig. 2. Lowest resolution size H and highest resolution size h in the interpolating wavelet method.

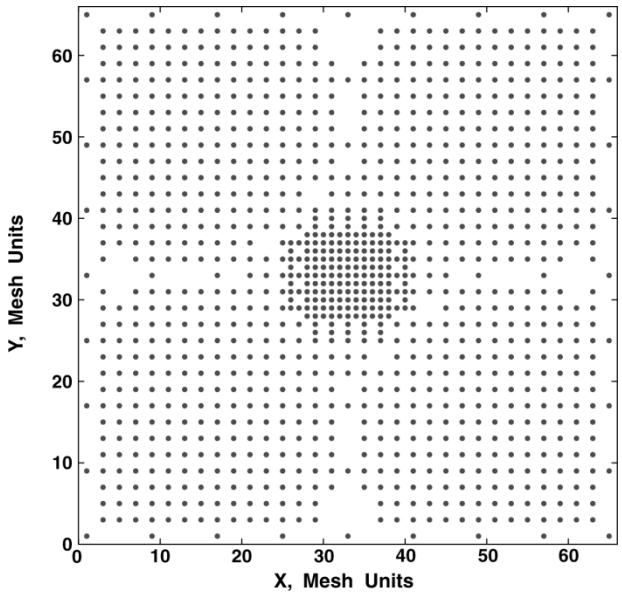


Fig. 3. Location in X–Y plane of nonzero coefficients of interpolating wavelet expansion, assuming error threshold of $\epsilon = 10^{-5}$.

Assume that we have a continuous function, which is the initial condition value. The construction of the mesh can be demonstrated for a rectangular domain as follows.

1) Set the smallest discretization value h and calculate all mesh points equally spaced by the distance h in the given domain, as illustrated in Fig. 2. This is the refinement limit.

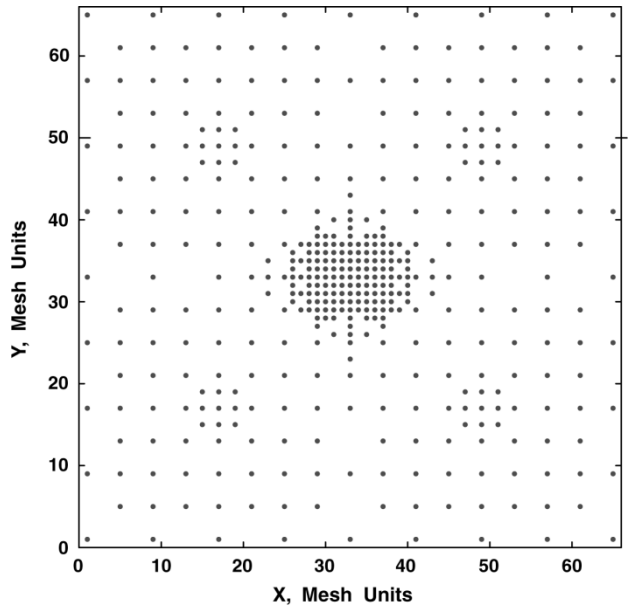


Fig. 4. Location in X–Y plane of nonzero coefficients of interpolating wavelet expansion, assuming error threshold of $\epsilon = 10^{-4}$.

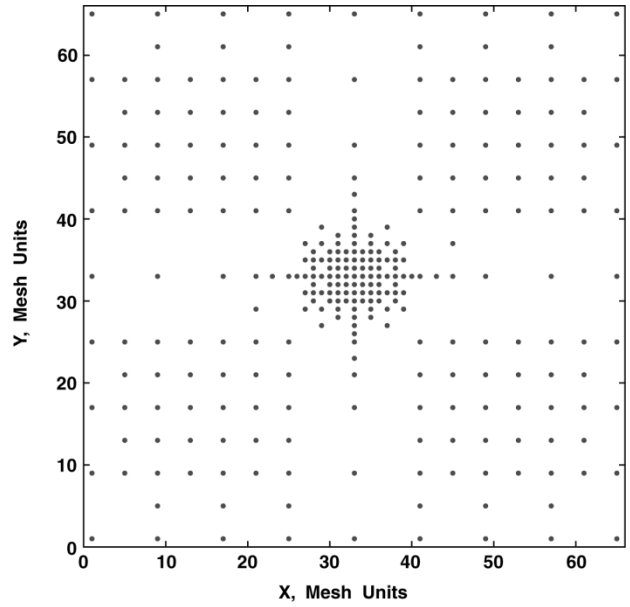


Fig. 5. Location in X–Y plane of nonzero coefficients of interpolating wavelet expansion, assuming error threshold of $\epsilon = 10^{-3}$.

TABLE III
NUMBER OF SIGNIFICANT COEFFICIENTS OF THE TEST FUNCTION FOR DIFFERENT VALUES OF THE THRESHOLD PARAMETER

Threshold parameter ϵ	The number of significant coefficients
0.0	4225
0.00001	1073
0.0001	429
0.001	291

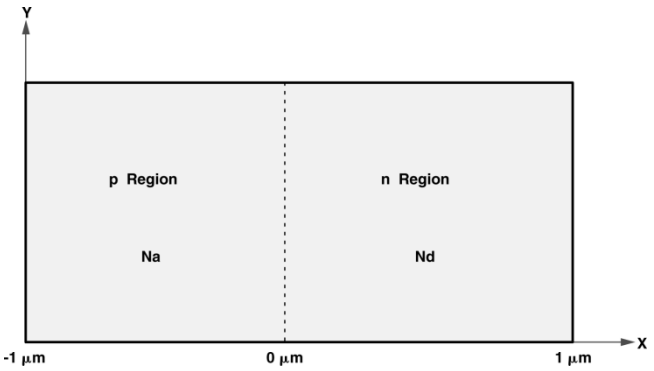


Fig. 6. Diagram of 1-D silicon p-n junction.

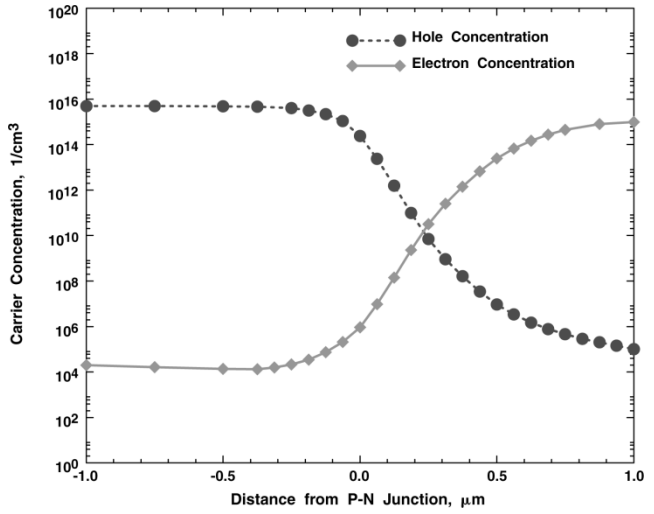


Fig. 7. Electron and hole carrier concentrations for 1-D abrupt silicon p-n junction at zero external bias.

2) Set the largest discretization value H and calculate all mesh points equally spaced by the distance H in the given domain. This is the coarsest limit. These mesh points will define interpolating scaling functions. Let us assign the initial function values to coarse mesh points. Function values corresponding to coarse mesh points will always be present in the SPR. In Fig. 2, these mesh points with function values are denoted by black dots. There is a limitation on the choice of H. The condition $H = 2^J h$ for a fix value of J must hold. J is then called an interpolating wavelet level.
3) For all intermediate mesh points equally spaced by the distance $H/2$ and different from points defined during the previous step, we calculate their interpolating values. We then compare the interpolated value with the real value of the function at this point. If the difference is less than the given threshold value ϵ, the corresponding node is excluded from consideration; otherwise, it is added to the SPR. Step 3 is repeated with the half-space distance until the refinement limit is reached.

As an example, we use a function that is smooth and slowly varying, except for a small region around the point with coordinates $(1/2, 1/2)$

$$u_0(x, y) = e^{-\alpha((x-1/2)^2+(y-1/2)^2)} - 0.2 \cdot \sin(2\pi x)\sin(2\pi y)$$

where the peak slope is controlled by the parameter α. At the conclusion of the above procedure, we will obtain pictures similar to Figs. 3–5, in which each black point refers to the grid

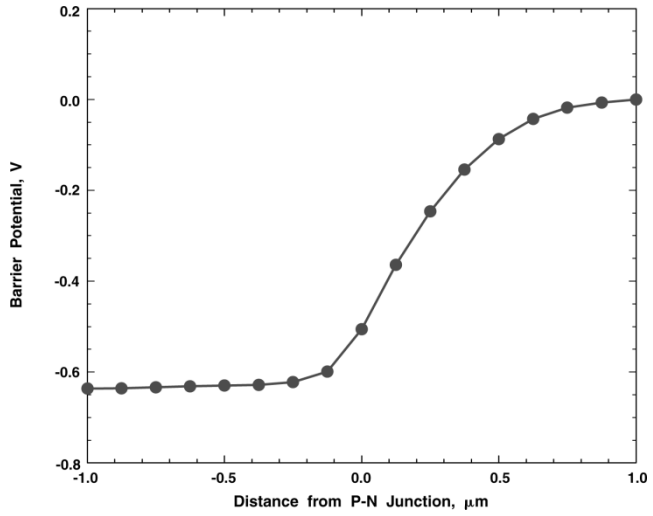


Fig. 8. Potential distribution for 1-D abrupt silicon p-n junction at zero external bias.

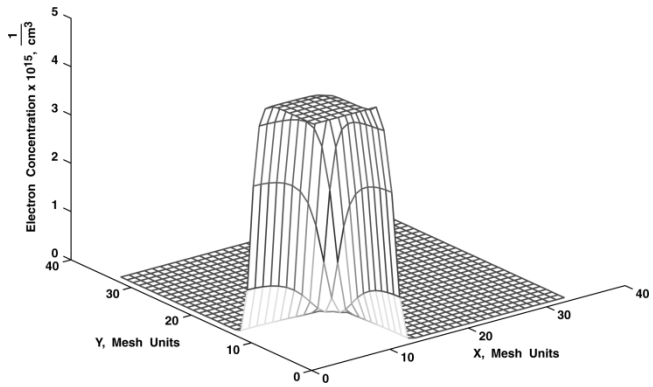


Fig. 9. Electron concentration for 2-D abrupt silicon p-n junction with zero external bias.

point with the assigned function value. These points correspond to significant coefficients of the test function for different values of the threshold parameter ϵ. Additional grid points are placed in the regions where sharp variations of the function occur.

Table III illustrates a variation in the number of significant coefficients of the test function versus the threshold parameter. The finest level of interpolating wavelets is $J = 3$. The smaller values of ϵ correspond to the finer mesh, until the refinement limit $\epsilon = 0$ is reached.

Theoretically, the different meshes may cause problems when we have, e.g., to add or multiply two different solution components. In such a situation, it is necessary to interpolate the missing points.

To account for possible changes in the solution during a time step, or to account for abrupt discontinuities or rapidly occurring transients in the nonlinear case, it is helpful to include the neighboring values, i.e., always to retain additional wavelet points in the SPR. After one or several time steps, we extend all the SPR to the complete solution on the finest mesh, and then using this mesh as an exact solution, we again form its SPR. If the solution changes in time rapidly, the new SPR will differ from the old one.

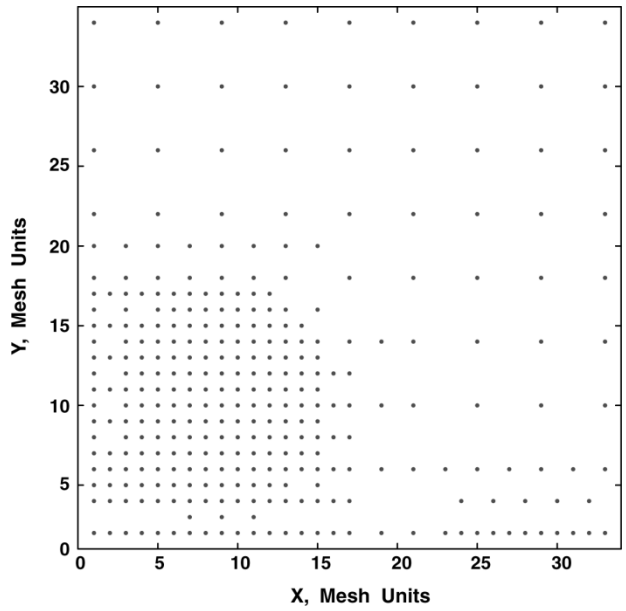


Fig. 10. Grid points of electron concentration for 2-D abrupt silicon p-n junction with zero external bias.

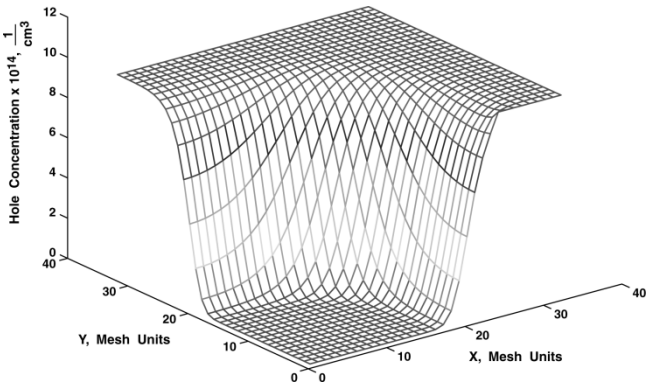


Fig. 11. Hole concentration for 2-D abrupt silicon p-n junction with zero external bias.

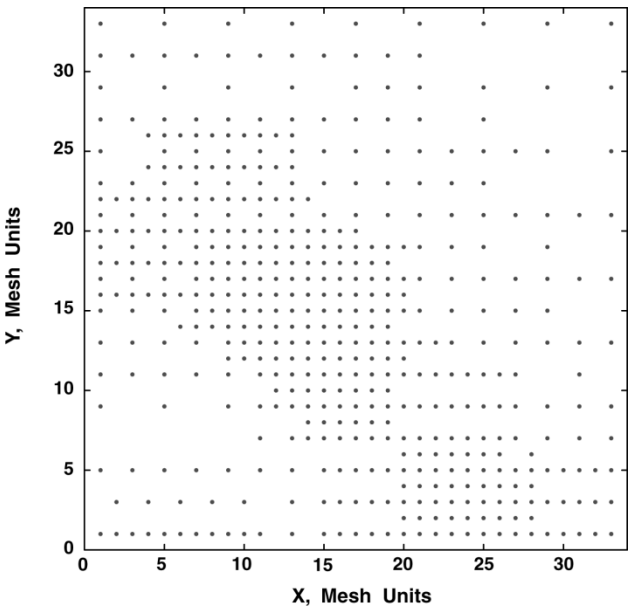


Fig. 12. Grid points of hole concentration in 2-D abrupt silicon p-n junction with zero external bias.

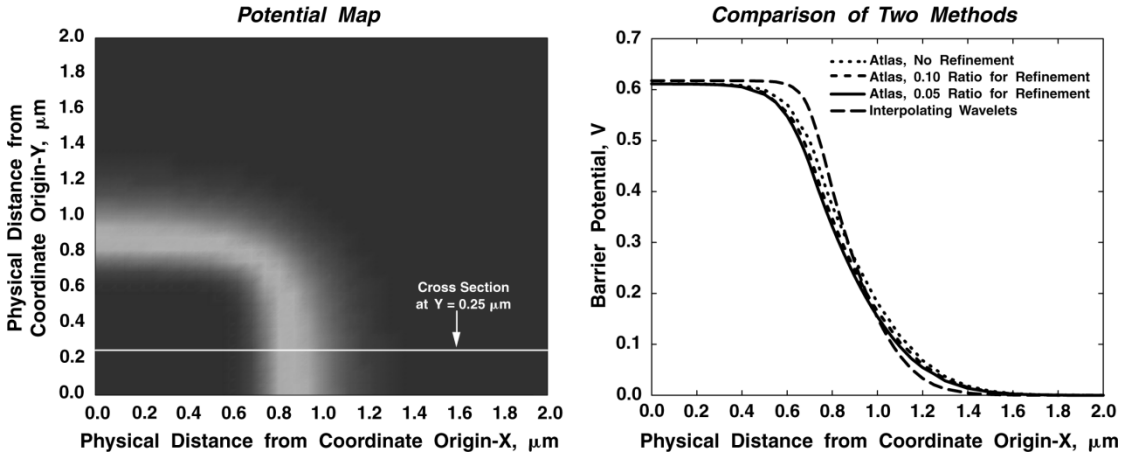


Fig. 13. Equipotential map and voltage profile for 2-D abrupt silicon p-n junction with zero external bias.

VII. Numerical Results

After a new algorithm is developed, it is always necessary to verify its correctness and to test its numerical accuracy. To this end, we have tested our computer programs with input parameters either from previously published papers or from well-known textbooks.

Example 1: Consider a 1-D silicon p-n junction in Fig. 6. The volume concentration of the implanted acceptors is $N_a = 5 \times 10^{15}$ cm^{-3} and the volume concentration of the implanted donors $N_b = 1 \times 10^{15}$ cm^{-3}. The 1-D problem has been discretized using the SPR with interpolating wavelets. Unlike analytic solutions, in which assumptions are made to simplify the mathematics, we have faithfully followed the tedious numerical procedures outlined in the previous sections.

The resulting electron and hole concentrations for an abrupt silicon p-n junction with zero external bias are presented in Fig. 7. The potential distribution appears in Fig. 8. Markers on the curves show corresponding mesh points. It is apparent from the figures that all components of the solution have their own meshes. In the case of the potential distribution, this mesh is quite coarse. Numerical calculations show that several hundred iterations (from 250 to 400) are required to achieve a steady-state solution of the equations. The number of iterations depends on the value of the threshold parameter ϵ. As has been noted before, the smaller parameter leads to the finer mesh and more iteration steps to achieve the converged solution. As expected, these two figures are in excellent agreement with [24, Figs. 2-2-2, 2-2-7], indicating the superior numerical precision of this new method.

Example 2: Consider an abrupt n$^+$-p diode in 2-D, as depicted in Fig. 1. The doping concentration under the left-hand-side contact is $N_d = 5.0 \times 10^{15}$ cm^{-3}. In the substrate $N_a = 1.0 \times 10^{15}$ cm^{-3} (p-type). Figs. 9 and 10 illustrate the distribution of the electron concentration and its corresponding mesh. The number of nodes in the mesh is 325. Figs. 11 and 12 illustrate the distribution of the hole concentration and its corresponding mesh. The number of nodes in the mesh is 613. It is apparent that the SPR of the electron concentration differs from the SPR of the hole concentration. Further, each component of the solution has its own mesh, which is optimally adapted to the behavior of that component. The 2-D electron and hole concentrations have been compared with those from the commercial package ATLAS. The two sets of results are compatible, though not exactly identical. To extend the investigation, we plotted the 2-D potential distribution of our results as an equipotential map in the left-hand-side panel of Fig. 13. A potential profile at $y = 0.25$ μm from the map was taken and appears as the right-hand-side panel of Fig. 13. Again, our results and those from ATLAS were in good agreement.

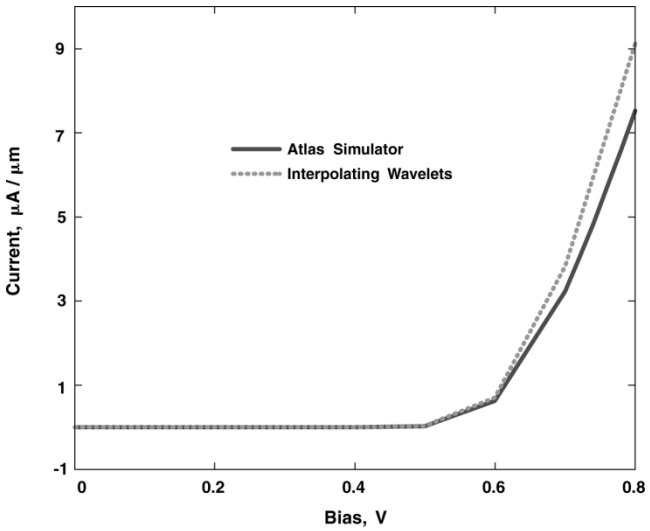


Fig. 14. Comparison of *I–V* curves between ATLAS and wavelet results for a 2-D abrupt silicon p-n junction.

It is often difficult to judge the precision of two numerical solutions when they show slight differences from one another. Detailed laboratory measurements would appear to be the only way to resolve these differences. However, it is extremely difficult to measure the potential profiles in such a tiny region inside the diode. The literature does document a few so-called "hero experiments," in which specially designed diodes have been fabricated and passivated, and then probed with a scanning tunneling microscope to create approximate measurements of the field distributions. However, it is believed by practitioners in this field that these measurements are sufficiently indirect such that the simulation results are probably closer representations

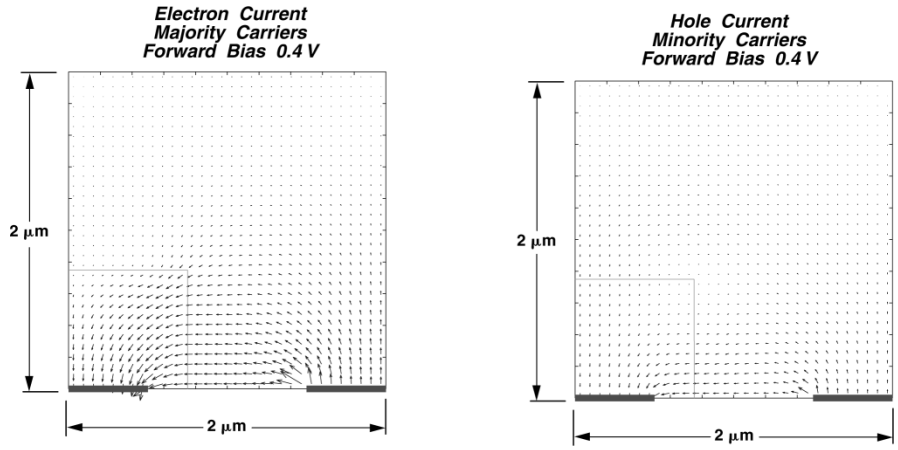


Fig. 15. Current distributions for idealized abrupt silicon p-n junction.

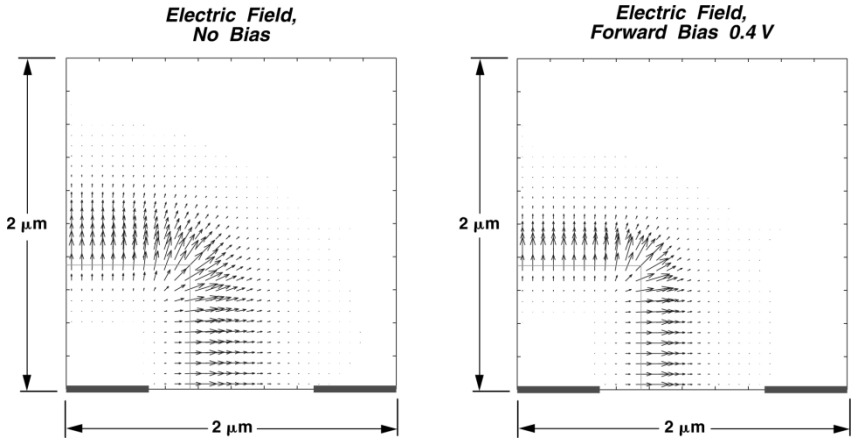


Fig. 16. Electric-field distributions for idealized abrupt silicon p-n junction.

of the actual device behavior than the reported measurements. Thus, we believe that we were justified in integrating the simulated current densities to obtain the device circuit parameter as the *I–V* curve, as depicted in Fig. 14. The two curves in Fig. 14, calculated with ATLAS and with our method, do exhibit small differences when the device bias voltage exceeds 0.6 V and the relative error reaches 19% at 0.8 V. The discrepancy between the two curves is due to slight differences in material parameters, including mobility, intrinsic density, etc. If we account for the exponential behavior of the *I–V* curves, the small discrepancy is quite satisfactory; in this case (unlike for the potential distribution discussed above), detailed laboratory tests might be able to resolve the differences between these two methods.

It is worth noting the numerical efficiency of the new approach. In the solution of the potential distribution, 423 nodes were needed to achieve a precision of 1.6% for the wavelets, while for a 5% precision, the Silvaco ATLAS simulator required 1756 triangles.

In Fig. 15, we have plotted the normalized electron current (the majority carrier current) and the hole current (the minority carrier current), both under a forward bias of 0.4 V. Fig. 16 illustrates the electric-field distributions for 0- and 0.4–V bias. It can be clearly seen that the depletion region shrinks as the forward bias is applied. The number of iterations required in the computation was approximately 1000. The meshes developed with the computationally efficient nonuniform wavelet method described in the previous paragraphs were generated with the threshold parameter $\epsilon = 0.1$ for the electron and hole concentrations, and $\epsilon = 0.01$ for the potential. The full mesh of the uniform grid consisted of 1089 nodes, in contrast to the 423 nodes for the nonuniform wavelet approach. The computational results associated with the nonuniform meshes compared favorably with simulation results obtained from the full mesh in its refinement limit of size h. The number of iterations in the numerical examples presented above were in the range of 1000–5000 for different values of the threshold parameter ϵ.

VIII. Conclusions

In this paper, we have applied interpolating wavelets to the modeling of both the transient and steady-state behavior of typical active microwave devices. The so-called SPR method was introduced to obtain a fully adaptive nonuniform mesh. Field, carrier, and current distributions inside a 2-D diode have been solved. These parameters can be used to study the interactions between an active device and its surrounding electromagnetic environment. Several numerical examples demonstrated good agreement with results from commercial simulation tools. The interpolating wavelets represent functions in a more efficient manner, so that many fewer nodes are required to achieve

high-precision results. The efficiency of this new approach will become even more profound when the modeling of these devices is extended to a full three dimensions.

ACKNOWLEDGMENT

The authors would like to thank Dr. J. Murphy, Defense Advanced Research Projects Agency/Microsystems Technology Office, Arlington, VA, and Dr. C. Hanson, Space and Warfare Systems Center Code 80, San Diego, CA, for administrative support and many helpful discussions, Dr. E. Daniel, Mayo Foundation, Rochester, MN, for technical assistance, and R. Techentin, S. Richardson, D. Jensen, E. Doherty, T. Funk, and L. Sievers, Mayo Foundation, Rochester, MN, for assistance in the preparation of text and figures.

REFERENCES

[1] C.-N. Kuo, B. Houshmand, and T. Itoh, "Full-wave analysis of packaged microwave circuits with active and nonlinear devices: A FDTD approach," *IEEE Trans. Microwave Theory Tech.*, vol. 44, pp. 819–829, May 1997.

[2] S. El-Ghazaly and T. Itoh, "Electromagnetic interfacing of semiconductor devices and circuits," in *IEEE MTT-S Int. Microwave Symp. Dig.*, June 1997, pp. 151–154.

[3] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM, 1992.

[4] G. Wang and G. Pan, "Full wave analysis of microstrip floating structures by wavelet expansion method," *IEEE Trans. Microwave Theory Tech.*, vol. 43, pp. 131–142, Jan. 1995.

[5] X. Zhu and G. Pan, "On application of fast and adaptive periodic Battle-Lemarie wavelets to modeling of multiple lossy transmission lines," *J. Comput. Phys.*, vol. 132, pp. 299–311, Apr. 1997.

[6] B. Jawerth and W. Sweldens, "An overview of wavelet based multiresolution analyses," *SIAM Rev.*, vol. 36, no. 3, pp. 377–412, Sept. 1994.

[7] M. Krumpholz, H. G. Winful, and L. P. B Katehi, "Nonlinear time domain modeling by multiresolution time domain (MRTD)," *IEEE Trans. Microwave Theory Tech.*, vol. 45, pp. 385–395, Mar. 1997.

[8] E. M. Tentzeris, R. L. Robertson, L. P. B. Katehi, and A. Cangellaris, "Space and time adaptive gridding using MRTD technique," in *IEEE MTT-S Int. Microwave Symp. Dig.*, 1997, pp. 337–340.

[9] L. Jameson, "On the wavelet optimized finite difference method," NASA Langley Res. Center, Hampton, VA, Tech. Rep. ICASE 94-9, Mar. 1994.

[10] G. Beylkin and J. Keiser, "On the adaptive numerical solution of nonlinear partial differential equations in wavelet bases," *J. Comput. Phys.*, vol. 132, pp. 233–259, 1997.

[11] M. Holmström, "Solving hyperbolic PDE's using interpolating wavelets," Uppsala Univ., Uppsala, Sweden, Tech. Rep. 189/1996, 1996.

[12] G. Deslauriers and S. Dubuc, "Symmetric iterative interpolation processes," *Construct. Approx.*, vol. 5, no. 1, pp. 49–68, 1989.

[13] D. L. Donoho, "Interpolating wavelet transforms," Dept. Statistics, Stanford Univ., Stanford, CA, preprint, 1992.

[14] W. Sweldens, "The lifting scheme: A construction of second generation wavelets," *Siam J. Math. Anal.*, vol. 29, no. 2, pp. 511–546, 1997.

[15] A. H. Marshak, "Carrier densities and emitter efficiency in degenerate materials with position dependent band structure," *Solid State Electron.*, vol. 21, pp. 429–434, 1978.

[16] A. De Mari, "An accurate numerical steady-state one dimensional solution of the p-n Junction," *Solid State Electron.*, vol. 11, pp. 33–58, 1968.

[17] P. Markowich, *The Stationary Semiconductor Device Equations*. Berlin, Germany: Springer-Verlag, 1985.

[18] D. L. Scharfetter and H. K. Gummel, "Large-signal analysis of a silicon read diode oscillator," *IEEE Trans. Electron Devices*, vol. ED-16, pp. 64–77, Jan. 1969.

[19] J. M. Ortega and W. G. Poole, *An Introduction to Numerical Methods for Differential Equations*. New York: Pitman, 1981.

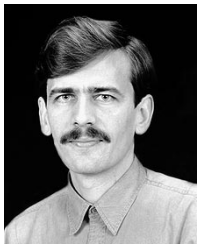
[20] H. A. van der Vorst, "Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems," *SIAM J. Sci. Statist. Comput.*, vol. 13, no. 2, pp. 631–644, Mar. 1992.

[21] M. S. Adler, "A method for terminating mesh lines in finite difference formulations of the semiconductor device equations," *Solid State Electron.*, vol. 23, pp. 845–853, 1980.

[22] A. F. Franz, G. A. Franz, S. Selberherr, C. Ringhofer, and P. Markovich, "Finite boxes—A generalization of the finite difference method suitable for semiconductor device simulation," *IEEE Trans. Electron Devices*, vol. ED-30, pp. 1070–1082, Sept. 1983.

[23] J. Bürgler, *Discretization and Grid Adaptation in Semiconductor Device Modeling*. Konstanz, Germany: Hartung-Gorre, 1990.

[24] M. Shur, *Physics of Semiconductor Devices*. Englewood Cliffs, NJ: Prentice-Hall, 1990, sec. 2-2.



Mikhail V. Toupikov (S'96–M'97) received the Candidate of Science degree in applied mathematics from the Moscow State University, Moscow, Russia, in 1987, and the Ph.D. degree in electrical engineering from Arizona State University, Tempe, in 1997.

From 1986 to 1993, he was a Scientific Researcher in the Department of Computational Mathematics and Cybernetics, Moscow State University. In 1994, he was a Visiting Scholar at the University of Wisconsin–Milwaukee. Since 1997, he has been a Post-Doctoral Faculty Research Associate at Arizona State University. His research interests are in the area of numerical methods in electromagnetics and mathematical modeling of microwave circuit devices.

Dr. Toupikov is a member of Eta Kappa Nu.

Guangwen (George) Pan (S'81–M'84–SM'94) received the M.S. and Ph.D. degrees from the University of Kansas, Lawrence, in 1982 and 1984, respectively, both in electrical engineering.

From 1984 to 1985, he was a Post-Doctoral Fellow at the University of Texas. In 1985, he joined the Mayo Foundation, where he was involved with the theory for electromagnetic modeling of packaging and interconnections for high-speed integrated circuits and printed circuit boards. From 1986 to 1988, he was an Associate Professor of electrical engineering at South Dakota State University. From 1988 to 1993, he was an Associate Professor in the Department of Electrical Engineering and Computer Science, University of Wisconsin–Milwaukee, with promotion to Professor in 1993. Since 1995, he has been Professor and Director of Electronic Packaging Laboratory, Department of Electrical Engineering, Arizona State University, Tempe. His interests are the modeling of high-frequency and wide-band signal-propagation phenomena in high-performance integrated circuits, multichip modules, and printed circuit boards, and in the design of high-performance antennas for a variety of applications.

Barry K. Gilbert (S'62–M'70–SM'87–F'98) received the B.S. degree in electrical engineering from Purdue University, West Lafayette, IN, and the Ph.D. degree in physiology and biophysics (with minors in electrical engineering and applied mathematics) from the University of Minnesota, Minneapolis–St. Paul.

He currently is a Staff Scientist in the Department of Physiology and Biophysics, Mayo Foundation, Rochester, MN, and the Director of the Special-Purpose Processor Development Group, Mayo Foundation (36 full-time staff members). His research interests include the development of algorithms for the real-time analysis of wide-bandwidth image and signal data, the design of specialized signal-processing computers to execute these tasks, the development of computer-aided design (CAD) tools to allow the timely design of high-complexity digital-signal processors, the advancement of high-performance integrated-circuit technologies, such as gallium arsenide and indium phosphide, which can be used to assemble very high-performance signal processors, and the development of advanced electronic packaging technologies such as multichip modules, which will be capable of supporting digital integrated circuit-based processors operating at gigahertz system clock rates.